Logarithms and log-transformations

Logarithms (frequently referred to as 'logs') are often used in statistics. Medical statisticians log-transform skewed data to make the distribution of the data more symmetrical and this helps data 'behave better' by meeting the assumptions of statistical models. When plotting graphs, log-transforming makes curved data fall on lines which are more straight, and 'flattens' data, drawing in extreme values which enables patterns in the data to be seen more clearly.

In my blog, when I refer to logs I will mean 'natural' logs for which the base is e, known as Euler's number, and it is equal to 2.718 to 3 decimal places. I will write the natural log of x as *ln(x)*. Medical statisticians typically use natural logs. In some sciences, logs to base 10, also known as the 'common logarithm' is more common. A bit of maths (see below if you're interested) shows us

$$ln(x) = log10(x) \times ln(10)$$

$$log10(x) = \ln(x) \times log10(e)$$

$$e = 10^{\frac{1}{\ln(10)}}$$

In computer programs and software packages, natural logs of x is written as *log(x)* in R and SAS, *LN(x)* in SPSS and EXCEL, and either *ln(x)* or *log(x)* in STATA. On calculators, the button to calculate the natural log of a number is *ln*. To antilog (cancel out) natural logs, we use the exponential function. This is written as *exp(x)* in R, SAS and STATA, and *EXP(x)* in SPSS and EXCEL. In calculators, it is given as e^x, and this may have its own button, as on a mobile phone app, or it may be accessed on a hand-held calculator by pressing the *SHIFT* and *ln* keys in sequence.

Where did the equations come from?

To calculate ln(x):		
Start with	$log10(x) = y$	**(equation 1)**
Which means that	$x = 10^y$	
Taking natural logs of both sides	$\ln(x) = \ln(10^y)$	
Using the 3rd law of logs $ln(x^n) = n \times ln(x)$	$\ln(x) = y \times \ln(10)$	
Substitute for y using equation 1	$\ln(x) = log10(x) \times ln(10)$	**(equation 2)**
To calculate log10(x):		
By a similar process, starting with	$ln(x) = y$	**(equation 3)**
Which means that	$x = e^y$	
Taking logs of base 10 of both sides and using 3rd law of logs	$log10(x) = y \times log10(e)$	
Substitute for y using equation 3	$log10(x) = \ln(x) \times log10(e)$	
To calculate e:		
Substitute for log10(x) in equation 2 and cancel out ln(x)	$1 = log10(e) \times ln(10)$	
Rearranging the equation	$log10(e) = \frac{1}{\ln(10)}$	
Which means that	$e = 10^{\frac{1}{\ln(10)}}$	

Dr Kathy Taylor teaches data extraction in Meta-analysis, https://www.conted.ox.ac.uk/courses/meta-analysis</link> This is a short course that is also available as part of our MSc in Evidence-Based Health Care

https://www.conted.ox.ac.uk/about/msc-in-evidence-based-health-care, MSc in Medical Statistics

https://www.conted.ox.ac.uk/about/msc-in-ebhc-medical-statistics, and MSc in Systematic Reviews

https://www.conted.ox.ac.uk/about/msc-in-ebhc-systematic-reviews

Follow updates on this blog and related news on Twitter @dataextips